

**Ciência da Computação – Universidade Federal do Agreste de  
Pernambuco (BCC-UFAPE) – Garanhuns – PE – Brasil**

**Projeto da Disciplina**

**Disciplina Optativa: Fundamentos de Ciência de Dados**

**Professores Responsáveis**

Ryan Azevedo & Assuero Ximenes  
{ryan.azevedo, assuero.ximenes}@ufape.edu.br

**Parte 2 – Jogadores NBA**

Todo o projeto deve ser implementado/desenvolvido em Python. A parte de visualização de Dados deve ser em Python ou em Python e *Java Script*, os gráficos podem ser gráficos dinâmicos ou estáticos. Podem usar funções do Python para implementar as estatísticas e os algoritmos necessários.

**Engenharia de Dados (Para Todo o Projeto)**

A primeira etapa do projeto consiste na Engenharia de Dados. As atividades que devem ser realizadas nessa etapa são:

- Extração dos dados dos jogos e jogadores (Apenas do seu time e dos seus Jogadores) da NBA temporada 23-24 (passada) e 24-25 (atual).
- Nas bases [isso pode ser e [deve ser] feito durante o ciclo de desenvolvimento] verificar dados ausentes, valores redundantes, ruidosos, inconsistentes, enviesados e valores outliers caso existam.
- Transformar os dados, verificar tipos, converter dados qualitativos para quantitativos, excluir ou não utilizar colunas do *Dataset* que não impactam [ainda] no projeto [Ex.: salário do jogador], normalizar valores numéricos quando necessário, preencher dados ausentes com dados reais e disponíveis pela entidade/empresa/corporação NBA [no caso de placar de jogos não extraídos automaticamente].
- Como *amostras de dados* vamos utilizar todos os dados do *Dataset* (100% dos dados que nos interessa). [Não estamos preocupados com desempenho computacional].
- O procedimento de separação [Treinamento e teste] do conjunto de dados será realizado apenas durante o desenvolvimento dos modelos preditivos e estatísticos que usaremos. Aqui usaremos técnicas automáticas para evitar o *data leakage*.

Com os dados coletados, passamos para a fase de desenvolvimento das funcionalidades relacionadas ao time/equipe.

## Funcionalidades/Requisitos – Jogadores<sup>1</sup>

O sistema deve proporcionar aos *stakeholders* um conjunto de funcionalidades que os auxiliem no processo de tomada de decisão. Os dados utilizados [para os jogadores] serão os da **atual temporada (24-25)** da NBA. O sistema deverá atender aos Requisitos Funcionais (RF) definidos e detalhados neste documento de especificação.

### Requisitos Funcionais:

**RF1** – Apresentar os seguintes dados dos Jogadores [Dependendo de “onde esteja extraíndo” os dados, existe um Id, exiba também o Id de cada Jogador] (Vide Tabela 1). Se o jogador for um *Rookie*, converter para o número 0.

Nome	Altura	Peso	Idade	Experiência	Posição	Universidade	Salário
------	--------	------	-------	-------------	---------	--------------	---------

Tabela 1. Dados dos Jogadores

**RF2** – Para cada jogador do seu time apresente os dados conforme Tabela 2. [OBS.: O sistema deve fornecer automaticamente os dados de todos os jogos ocorridos durante a temporada atual em tempo real].

Data do Jogo	Adversário	V ou D <sup>2</sup>	Casa/Fora <sup>3</sup>	PTS <sup>4</sup>	REB	AST	Placar do Jogo
--------------	------------	---------------------	------------------------	------------------	-----	-----	----------------

Tentativas de Cestas de 3 PTS	Cestas de 3 PTS Marcados <sup>5</sup>	Tempo de Permanência do Jogador em Quadra
-------------------------------	---------------------------------------	---

Tabela 2. Dados das Partidas

**RF3** – O sistema deve fornecer os dados das partidas que o usuário escolher. Ex. Partidas contra time X e apresentar esses dados conforme Tabela 3.

Data do Jogo	Adversário	V ou D <sup>6</sup>	Casa/Fora <sup>7</sup>	PTS <sup>8</sup>	REB	AST	Placar do Jogo
--------------	------------	---------------------	------------------------	------------------	-----	-----	----------------

Tentativas de Cestas de 3 PTS	Cestas de 3 PTS Marcados <sup>9</sup>	Tempo de Permanência do Jogador em Quadra
-------------------------------	---------------------------------------	---

Tabela 3. Dados das Partidas Específicas

---

<sup>1</sup> Todos os alunos já sabem seus jogadores e podem verificar nos Apêndice 1 e 2 deste documento.

<sup>2</sup> V – Vitória | D -Derrota

<sup>3</sup> Jogou fora ou dentro de casa

<sup>4</sup> PTS – Pontos | REB – Rebotes | AST – Assistências

<sup>5</sup> Quantidade de cestas marcadas

<sup>6</sup> V – Vitória | D -Derrota

<sup>7</sup> Jogou fora ou dentro de casa

<sup>8</sup> PTS – Pontos | REB – Rebotes | AST – Assistências

<sup>9</sup> Quantidade de cestas marcadas

**[Início - Apenas dos jogadores definidos e referentes e ao seu time]**

**RF4** – O sistema deve apresentar a quantidade de jogos realizados dentro e fora de casa e a quantidade de jogos dentro e fora de casa contra um determinado time [da escolha do usuário].

**RF5** – Apresentar e calcular a **média** de pontos, rebotes e assistências dos jogadores

- **RF5 – A** – Apresentar ao usuário a porcentagem de pontos, rebotes e assistências abaixo da média

**RF6** – Apresentar e calcular a **mediana** de pontos, rebotes e assistências dos jogadores.

- **RF6 – A** – Apresentar ao usuário a porcentagem de pontos, rebotes e assistências abaixo da mediana

**RF7** – Apresentar e calcular a **moda** de pontos, rebotes e assistências dos jogadores. Exibir a quantidade de vezes que a moda aparece para cada item [pontos, rebotes e assistências].

- **RF7 – A** – Apresentar ao usuário a porcentagem de pontos, rebotes e assistências abaixo da média

**RF8** – Apresentar o Desvio Padrão de pontos, rebotes e assistências dos jogadores. Quanto mais próximo de zero, mais agrupado em torno da média os dados estão.

**RF9** – O sistema deve apresentar a quantidade de pontos, assistências e rebotes de toda a carreira do jogador.

**RF10** – O sistema deve apresentar a quantidade de pontos, assistências e rebotes de toda a carreira <sup>10</sup>do jogador e comparada com a atual temporada conforme Tabela 4.

Estatísticas	Total de jogos	Média de Pontos	Média de Assistências	Média de Rebotes	Minutos em Quadra
Temporada Atual					
Carreira					

Tabela 4. Dados Comparativos

**RF10** – Apresentar gráficos de desempenho dos seus jogadores [**temporada atual**] para compor o *Dashboard* do projeto:

- Gráfico de distribuição de pontos por jogo em relação a média, mediana e moda
- Gráfico de distribuição de rebotes por jogo em relação a média, mediana e moda
- Gráfico de distribuição de assistências por jogo em relação a média, mediana e moda
- Gráfico **BOX PLOT** de pontos, rebotes e assistências por jogo [detalhar todas as informações (Max, Min, Mediana, *Outliers*, etc.) possíveis nos quartis]

**[Fim – Apenas dos jogadores definidos e referentes e ao seu time]**

---

<sup>10</sup> Caso ele seja um *Rookie* não inserir dados da carreira

**RF11** – Para cada tabela e/ou requisito que exiba dados gerados pelo sistema será necessário que os dados sejam salvos em um arquivo *c.v.*

**RF12** – Para cada gráfico gerado pelo sistema será necessário que sejam exibidos em formato HTML e abertos no *Browser*.